# A Smart Visual Assistance System For Visually Impaired People In Real-World Environments

Harinakshi C
*Assistant Professor, Department of ISE*
*Sahyadri College of Engineering and Management*
Mangalore, Karnataka, India
harinidece25@gmail.com

Nehal Hosalikar
*Department of ISE*
*Sahyadri College of Engineering and Management*
Mangalore, Karnataka, India
nhosalikar@gmail.com

Karthik Shetty
*Department of ISE*
*Sahyadri College of Engineering and Management*
Mangalore, Karnataka, India
shettykarthik1412@gmail.com

Keerthan K
*Department of ISE*
*Sahyadri College of Engineering and Management*
Mangalore, Karnataka, India
keerthankarkera101@gmail.com

Shetty Rachana Ravindra
*Department of ISE*
*Sahyadri College of Engineering and Management*
Mangalore, Karnataka, India
rachanashetty.2702@gmail.com

2024 International Conference on Intelligent Systems for Cybersecurity (ISCS) | 979-8-3503-7523-7/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCS61804.2024.10581293

***Abstract*—In a world emphasizing self-sufficiency and inclusivity, the visually impaired community faces significant challenges in everyday tasks, particularly handling currency and accessing information while shopping. According to data from the World Health Organization (WHO), the global population of visually impaired individuals is steadily increasing. Currently, there are approximately 285 million visually impaired individuals, with 39 million of them experiencing complete blindness and the remaining 217 million dealing with varying degrees of low vision. To address these obstacles, the project introduces a wearable assistive device featuring camera-on-glasses technology. At its core, this device includes a Currency Note Identification system that empowers visually impaired users to accurately identify currency denominations using MobileNet artificial neural network (ANN). Furthermore, it integrates Optical Character Recognition (OCR) technology for reading printed text . The system captures images, wirelessly transmits them to a dedicated mobile application, and delivers real-time auditory feedback. The design emphasises affordability and user accessibility, aiming to enrich the lives of visually impaired individuals by providing confidence and independence in financial management and shopping experiences. This innovative wearable device exemplifies the transformative potential of technology to promote inclusivity and equitable access.**

***Index Terms*—Smart-glasses, Visually Impaired, Mobile-app, Camera, Currency detection, OCR, MobileNet ANN.**

## I. INTRODUCTION

In a world that prizes self-sufficiency and inclusivity, the visually impaired community often faces significant hurdles when it comes to everyday tasks, particularly when handling currency and accessing information while shopping. In response to these pressing challenges, the project embarks on a mission to create a groundbreaking solution: a wearable assistive device [15] featuring camera-on-glasses technology [19], specifically designed to empower visually impaired [18] individuals in their shopping experiences. At the core of the project is the creation of a Currency Note Identification system, a dependable tool that allows visually impaired users to accurately identify the denomination of currency notes by using MobilNet ANN [5, 6] to accurately detect and identify currency denominations. Leveraging cutting-edge Optical Character Recognition (OCR) technology [3], the system processes a wide array of printed text [13], enabling users to read product names, aisle names, and even bills during their shopping experiences. The project unites meticulously crafted hardware and software components to seamlessly enhance the user experience. The camera-on-glasses captures images of currency, which are then transmitted to a dedicated mobile application [6] residing on the user's smartphone. Inside the app, advanced processing takes place, and the results are relayed directly to the user's audio output, providing real-time auditory feedback [10] about the currency in hand or the text on products and bills. Affordability is a key principle guiding the design process, with the aim of making this assistive technology accessible to a broader spectrum of users. Simultaneously, user accessibility remains at the forefront of the project's development. In essence, the system represents not just a technological innovation, but a commitment to empowering visually impaired individuals [16] with the confidence and independence to manage their finances and enjoy shopping experiences with the same ease and inclusivity as their sighted peers. The motivation behind articulating the project stems from a deep-seated commitment to addressing the pressing challenges faced by the visually impaired community in their daily lives[15].

## II. LITERATURE SURVEY

The proposed system in mentioned paper can identify indoor and outdoor objects, notify the users, and send all information to a remote server repeatedly at a fixed time interval proposed by Md. Atikur Rahman, Muhammad Sheikh Sadi [1]. The system uses a singleshot detector (SSD) model with MobileNet and TensorFlow-Lite to recognize objects in real time in both indoor and outdoor environments. The system also uses a voice assistant to provide audio feedback to the user about the objects in their surroundings. The system was evaluated on a dataset of over 100,000 images and achieved an accuracy of over 90% in recognizing a variety of objects, including people, animals, vehicles, and furniture.

Huiying Shen and James M. Coughlan [2] proposed a real-time system for finding and reading signs for visually impaired users. The system uses a smartphone camera to capture video images of the user's surroundings. A text detection algorithm is then used to identify signs in the video images. Once a sign is detected, the system uses an optical character recognition (OCR) algorithm to read the text on the sign and speak it aloud to the user.

Christos Liambas and Miltiadis Saratzidis [3] with the main idea of development of an autonomous mobile system for dictating text documents via image processing algorithm for visually impaired people proposed an Autonomous Optical Character Recognition (OCR) Dictating System designed to empower blind users in accessing printed text independently. The system incorporates OCR technology to recognize text from various sources and employs a dictation mechanism to convert the recognized text into audible speech, offering real-time assistance. The study not only focuses on the technical implementation of OCR but also emphasises the autonomous and user-centric aspects, catering to the specific needs of the blind community.

A deep learning-based method for identification of denominations of Indian Currency Rupee notes from their color images is presented by Shubham Mittal and Shiva Mittal [4]. Mainly focuses on a specific yet crucial aspect of currency recognition, addressing the context of Indian banknotes. The authors propose a Convolutional Neural Network (CNN)-based approach to accurately identify and classify Indian banknotes, catering to the unique design and features of the currency. The paper provides a detailed exploration of the CNN architecture employed, emphasising its effectiveness in handling the complexities associated with Indian banknotes.

Focusing on a situation when the ANN is already deployed on a desktop computer. The paper proposed by Milan Kostak, Ales Berger, et al [5]. explains the process of migration of an artificial neural network (ANN) to a smartphone device. The authors delve into the challenges and opportunities associated with adapting complex neural network models for execution on resource-constrained mobile devices. The study explores techniques and methodologies for optimising and enhancing the efficiency of ANNs when deployed on smartphones. By addressing the migration of neural networks to a mobile platform, the research contributes to the field of mobile computing and artificial intelligence, offering insights into the practical considerations and performance implications of running sophisticated models on smartphones.

The mentioned papers propose various innovative systems aimed at assisting visually impaired individuals in navigating their surroundings and accessing printed text independently. These technologies range from real-time object recognition systems using deep learning models [7,8] to OCR-based solutions for reading printed text. Overall, these contributions showcase the potential of technology to improve accessibility and independence for visually impaired individuals through innovative solutions. Through this integration of prior work, the aim is to leverage the advancements in assistive technology to enhance the functionality and effectiveness of the project in aiding visually impaired individuals.

## III. METHODOLOGY

The methodology presented in this paper outlines the development of a wearable assistive device, integrating camera-on-glasses technology, to address challenges faced by visually impaired individuals. Users initiate the process by facing the object of interest and activating a button on the glasses, capturing images transmitted wirelessly to a companion mobile app on the user's smartphone. The app processes the images and delivers auditory feedback, promoting independence for visually impaired individuals [20]. Smartphone connectivity to an audio output device ensures convenient reception of output.

The system consists of a Microcontroller, Camera Module, and Radio Frequency (RF) Transmitter, all seamlessly integrated into eyeglasses for user convenience. Additionally, a dedicated mobile application equipped with trained Machine Learning models is installed on the user's smartphone [5, 9]. The eye-glasses are connected wirelessly to the user's smartphone while the mobile application is running on it. The mobile application analyses the image provided by the eye-glasses and produces an audio output through the connected audio device. The architecture of the proposed system is shown in Fig.1.

### A. *Image Capturing Process*

The image capturing process in the proposed wearable assistive device is facilitated by the integration of the ESP32-CAM [17] microcontroller and the Arducam OV2640 camera module. Upon user prompt, the microcontroller embedded within the eyeglasses activates the camera module, initiating the capture of a real-time image of the user's surroundings. Leveraging the capabilities of the ESP32-CAM microcontroller and the high-resolution Arducam OV2640 camera, the

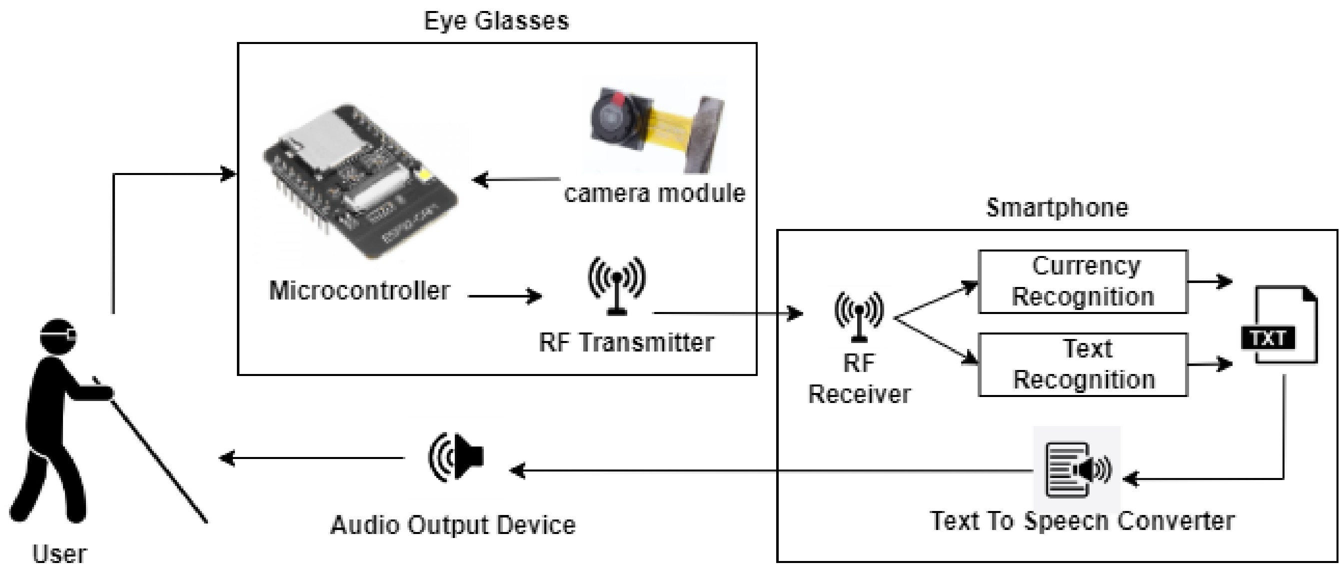


Fig. 1: Architecture Diagram of a Smart Visual Assistance System

device ensures precise and detailed image acquisition, crucial for accurate analysis and recognition tasks.

### *B. Wireless Data Transmission*

After capturing the image, the ESP32-CAM microcontroller [17] initiates the wireless transmission of the captured image data directly to the user's smartphone. Leveraging the integrated WiFi functionality of the ESP32 module, the device establishes a direct connection with the companion mobile application. This wireless communication mechanism enables seamless and real-time data transfer between the eyeglasses and the smartphone. By utilizing WiFi technology, the device ensures reliable and efficient transmission of image data, facilitating prompt analysis and response within the mobile application.

### *C. Data Processing and Analysis*

On receiving the transmitted image data, the mobile application on the user's smartphone begins the data processing and analysis stage [4]. The received image is forwarded as input to the pre-trained machine learning models [11] integrated within the mobile application. Leveraging the computational power of the smartphone, the application executes sophisticated data processing techniques to interpret the captured image accurately [5,9].

Currency recognition module plays a pivotal role in the functionality of the proposed wearable assistive device. The device employs sophisticated machine learning algorithm to accurately identify and classify various denominations of currency notes in real-time. The MobileNet model [5], pre-trained on the ImageNet dataset and fine-tuned for currency recognition, demonstrates robust performance in discerning subtle visual cues and features unique to different currency denominations. The code segment begins by loading a quantized TensorFlow Lite (TFLite) model [5], specifically trained for currency recognition, onto an interpreter object. This model, integrated with a convolutional neural network (CNN) architecture [4, 11], is capable of accurately identifying various denominations of currency notes. Subsequently, class labels corresponding to the recognized currency denominations are loaded from a text file. A custom image, representative of a currency note, is then preprocessed by resizing it to match the input dimensions expected by the model and normalizing its pixel values to a range between 0 and 1. The preprocessed image is fed into the interpreter as input, and inference is performed by invoking the interpreter. Post-inference, the output tensor containing the model's predictions is retrieved, and the index of the predicted class with the highest probability is determined. This index is used to fetch the corresponding class label from the loaded labels list. Finally, the predicted class label is predicted, providing insight into the denomination of the currency note depicted in the custom image.

Text recognition is a fundamental component of the proposed assistive technology, facilitated through the implementation of Optical Character Recognition (OCR) [3]. The integration of OCR functionality within the Flutter framework, leverages the FlutterTesseractOcr [14] package for cross-platform compatibility. Upon receiving an input image path, the OCR engine processes the image and extracts textual content using predefined parameters, such as the Page Segmentation Mode (PSM) and interword spacing preservation. In cases where no text is detected within the image, the system gracefully handles such scenarios by returning an informative message.

### *D. Hardware Implementation*

In the initial stages of the project development, SolidWorks was employed to create a comprehensive 3D model that served as the blueprint for the hardware implementation. Leveraging the powerful features of SolidWorks, the physical enclosure and mounting arrangements for the ESP32 S-CAM-CH340 Development Test Board, ESP32-CAM-MB MICRO USB Download Module, and a battery was meticulously designed. This 3D model enabled to visualize the spatial layout of components, ensuring optimal placement for functionality and accessibility.

Fig. 2: 3D Model of the Smart Glasses

The core of the setup comprised the ESP32 S-CAM-CH340 Development Test Board, a versatile WiFi and Bluetooth module equipped with an OV2640 camera. This module served as the primary microcontroller for the system, providing the necessary processing power and connectivity capabilities for image capture and analysis [1]. Additionally, the ESP32-CAM-MB MICRO USB Download Module was employed, which facilitated easy programming and debugging of the ESP32 CAM Development Board. With the microcontroller integrated alongside a dedicated photo-capture switch and powered by a battery mounted discreetly on the glasses, a successful implementation of a robust and efficient system was done.

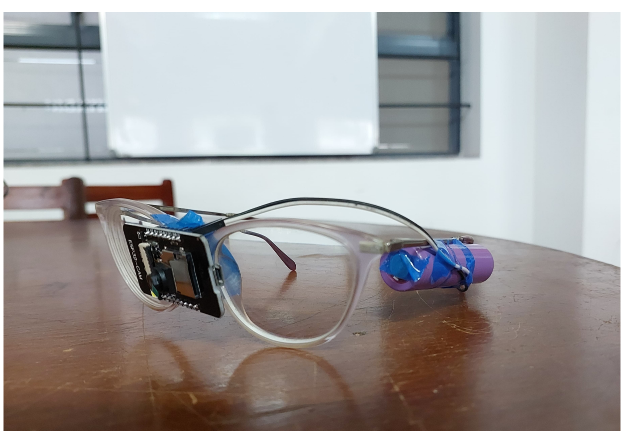

Fig. 3: Hardware Implementation

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

The experimentation framework was meticulously tailored, aligning with specific software and hardware interfaces to ensure optimal performance. For software development, Android Studio 2022.3.1 [9] and Visual Studio Code 1.82.1 was utilised as the primary integrated development environments (IDEs), coupled with Flutter 3.13.8 and PlatformIO 6.1.6 frameworks. The application was designed to accept images as input, which were then processed using the deployed models for currency recognition and text recognition. Additionally, Google Colab and Google Drive played crucial roles in facilitating collaboration and version control during the development of the machine learning models. On the hardware front, the setup featured an Intel Core i5-10300H processor, a minimum of 8GB RAM, and a hard disk capacity of at least 1TB to support the development environment and model deployment.

In the currency recognition module, the process begins with the capture of an image, which is then sent to the mobile application for analysis. The user interface (UI) of the application, as depicted in the accompanying Fig. 4, serves as the platform for displaying the results of this analysis. Here, the input image is showcased alongside the output, indicating which currency note is present in the image. Additionally, the UI provides insight into the confidence level of the prediction, offering users a measure of certainty regarding the accuracy of the model's classification.

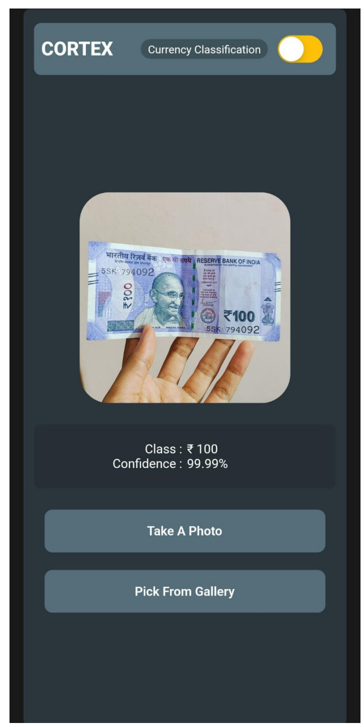


Fig. 4: Currency Recognition

In the text recognition module, the initial step involves capturing an image, which is subsequently transmitted to the application for further processing. The user interface (UI) depicted in the accompanying Fig. 5 serves as the platform for presenting the outcomes of this OCR-based processing. Within this UI, the input image is presented alongside the resulting text extracted from the item depicted in the image.

TABLE I: Performance of Currency Recognition Model

| Class (Notes) | Total Samples | Correctly Recognized | Accuracy (%) | Error Rate (%) | Precision | Recall | F1 Score |
|---|---|---|---|---|---|---|---|
| 10 | 215 | 214 | 99.94 | 0.05 | 100.00 | 99.53 | 99.76 |
| 20 | 276 | 276 | 99.88 | 0.12 | 99.28 | 100.00 | 99.64 |
| 50 | 272 | 268 | 99.65 | 0.35 | 99.26 | 98.53 | 98.89 |
| 100 | 301 | 301 | 99.83 | 0.17 | 99.01 | 100.00 | 99.50 |
| 200 | 205 | 205 | 99.59 | 0.40 | 96.69 | 100.00 | 98.32 |
| 500 | 223 | 222 | 99.94 | 0.05 | 100.00 | 99.55 | 99.77 |
| 2000 | 239 | 231 | 99.54 | 0.46 | 100.00 | 96.65 | 98.29 |

This integrated framework of image capture and OCR-based text extraction within the mobile application streamlines the process of text recognition, facilitating rapid interpretation of textual content from diverse captured items [9].

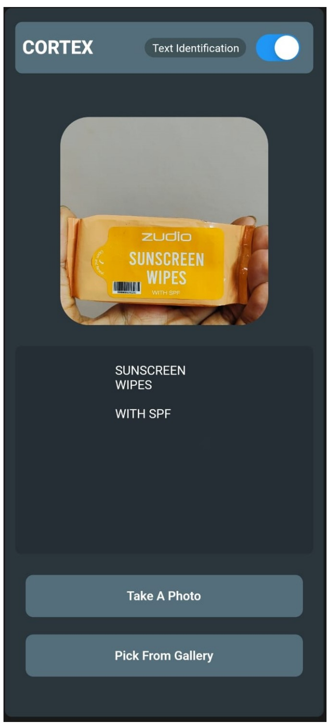


Fig. 5: Text Recognition

### A. Test Results

Confusion Matrix for the proposed model is shown in Fig.6. From the figure it can be observed that while 214 10-rupee notes, 276 20-rupee notes, 268 50-rupee notes, 301 100-rupee notes, 205 200-rupee notes, 222 500-rupee notes, 231 2000-rupee notes were classified accurately, 1 10-rupee note, 0 20-rupee note, 4 50-rupee notes, 0 100-rupee note, 0 200-rupee note, 1 500-rupee note, 8 2000-rupee notes were misclassified using the model. Thus, the correct classification rates of 10-rupee notes are 99.53%, 20-rupee notes are 100%, 50-rupee notes are 98.52%, 100-rupee notes are 100%, 200-rupee notes are 100%, while 99.55% for 500-rupee notes and 96.65% for 2000-rupee notes.

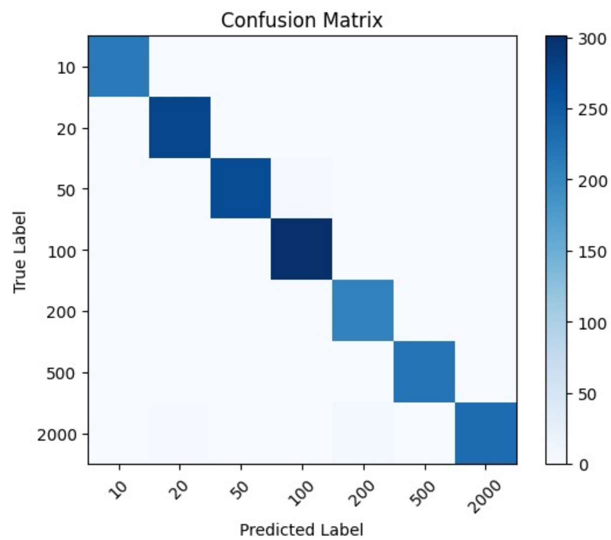


Fig. 6: Confusion Matrix

The Performance of the model is shown in Table I. Seven different types of note have been tested. An average of 1717 samples of objects is correctly recognized out of 1731 samples. The average accuracy is 99.76%. The highest accuracy is found to recognize the 10-rupee and 500-rupee currency notes which is 99.94%. The lowest accuracy is found to recognize the 2000-rupee currency note which is 99.54%.

TABLE II: Overall Performance Metrics for Currency Recognition Model

| Performance Metrics | Values (%) |
|---|---|
| Accuracy | 99.19 |
| Error Rate | 0.81 |
| Precision | 99.20 |
| Recall | 99.19 |
| F1 Score | 99.19 |

Overall performance metrics for the model is given in Table II. With an accuracy of 99.19% and an error rate of 0.81%, the model exhibits high correctness in identifying currency notes, with a precision of 99.20% effectively reducing false positive outcomes, while achieving a recall of 99.19% to capture almost all genuine currency notes, resulting in an F1 score of 99.19%, showcasing a balanced performance in currency recognition.

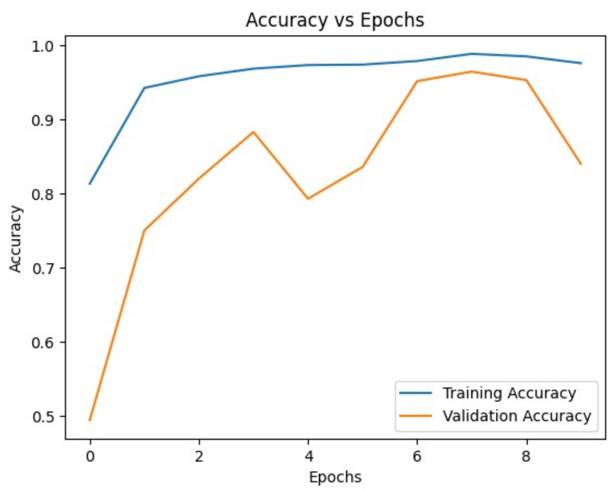


Fig. 7: The Accuracy Graph of the Developed Model

The graph in Fig. 7 shows the Training accuracy and Validation accuracy against each progressive epoch during training. To optimize training, cross-validation and early stopping was employed. Cross-validation [12] splits the dataset into subsets for iterative training and evaluation, guarding against overfitting and ensuring reliable performance estimates. Early stopping monitors validation set performance during training, halting when degradation is detected, saving time and preserving model generalization. These techniques enhance training efficiency and model reliability, and prevent overfitting.

## V. CONCLUSION

In conclusion, the development of a wearable visual aid [15] for the visually impaired represents a significant leap in enhancing the daily lives of individuals facing visual challenges. This project, with a focus on currency note identification and text reading, has been pursued with a commitment to user accessibility and feedback. The integration of advanced technologies, such as optical character recognition (OCR) and text-to-speech, has paved the way for a cost-effective, user-friendly solution. While the device does require the user to initiate the photo capture, its potential impact on the lives of the visually impaired remains substantial. By facilitating currency note identification and text reading, it offers a means for visually impaired individuals to engage more actively in daily activities, manage their finances, and access printed information independently.

## REFERENCES


[1] Md. A. Rahman and M. S. Sadi, "IoT Enabled Automated Object Recognition for the Visually Impaired," Computer Methods and Programs in Biomedicine Update, vol. 1, p. 100015, May 2021, ISSN: 2666-9900.

[2] Shen, H., Coughlan, J.M., "Towards a Real-Time System for Finding and Reading Signs for Visually Impaired Users", Computers Helping People with Special Needs, Lecture Notes in Computer Science, vol 7383, 2012.

[3] C. Liambas and M. Saratzidis, "Autonomous OCR dictating system for blind people", IEEE Global Humanitarian Technology Conference (GHTC), pp. 172-179, Seattle, WA, USA, 2016.

[4] S. Mittal and S. Mittal, "Indian Banknote Recognition using Convolutional Neural Network", 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU), pp. 1-6, Bhimtal, India, 2018.

[5] Kostak, M., Berger, A., Slaby, A., "Migration of Artificial Neural Networks to Smartphones", Lecture Notes in Computer Science, vol 12249, Springer, Cham, 2020.

[6] V. Veeramsetty, G. Singal, and T. Badal, "Coinnet: platform independent application to recognize Indian currency notes using deep learning techniques", Multimedia Tools and Applications, vol. 79, no. 31–32,pp. 22569–22594, May 2020.

[7] D. G. Lowe, "Object recognition from local scale-invariant features", Seventh IEEE International Conference on Computer Vision, vol.2., pp. 1150-1157, Kerkyra, Greece, 1999.

[8] Liu, W. et al., "SSD: Single Shot MultiBox Detector", Computer Vision – ECCV 2016, Lecture Notes in Computer Science, vol 9905, Springer, Cham, 2016.

[9] A. Mathur, A. Pathare, P. Sharma and S. Oak, "AI based Reading System for Blind using OCR", 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 39-42, Coimbatore, India, 2019.

[10] Rajalakshmi Pushparaman, "Design of Visual Aid for Blind People", Advances in Intelligent Systems and Computing, vol 394. Springer, New Delhi, 2016.

[11] Zhang, Q., Yan, W.Q. & Kankanhalli, "M. Overview of currency recognition using deep learning", Journal of Banking and Financial Technology 3, 59–69, 2019.

[12] Vishnu R and B. Omman, "Principal features for Indian currency recognition", IEEE India Conference (INDICON), pp. 1-8, Pune, India, 2014.

[13] S. Deshpande and R. Shriram, "Real time text detection and recognition on hand held objects to assist blind people", International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), pp. 1020-1024, Pune, India, 2016.

[14] R. Ani, E. Maria, J. J. Joyce, V. Sakkaravarthy and M. A. Raja, "Smart Specs: Voice assisted text reading system for visually impaired persons using TTS method", International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT), pp. 1-6, Coimbatore, India, 2017.

[15] Ruxandra Tapu, Bogdan Mocanu, Titus Zaharia, "Wearable assistive devices for visually impaired: A state of the art survey", Pattern Recognition Letters, Volume 137, pp. 37-52, ISSN 0167-8655, 2020.

[16] M. Singh, J. Chauhan, M. S. Kanroo, S. Verma and P. Goyal, "IPCRF: An End-to-End Indian Paper Currency Recognition Framework for Blind and Visually Impaired People", IEEE Access, vol. 10, pp. 90726-90744, 2022.

[17] M. J. Júnior, O. B. Maia, H. Oliveira, E. Souto and R. Barreto, "Assistive Technology through Internet of Things and Edge Computing", IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin), pp. 330-332, 2019.

[18] I. Joe Louis Paul, S. Sasirekha, S. Mohanavalli, C. Jayashree, P. Moohana Priya and K. Monika, "Smart Eye for Visually Impaired-An aid to help the blind people", International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1-5, Chennai, India, 2019.

[19] C. Yi, Y. Tian and A. Arditi, "Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons", IEEE/ASME Transactions on Mechatronics, vol. 19, no. 3, pp. 808-817, June 2014.

[20] M. A. Khan, P. Paul, M. Rashid, M. Hossain and M. A. R. Ahad, "An AI-Based Visual Aid With Integrated Reading Assistant for the Completely Blind", IEEE Transactions on Human-Machine Systems, vol. 50, no. 6, pp. 507-517, Dec. 2020.